

APPENDIX D
WORKSHOP CHARGE

Workshop on Selecting Input Distributions for Probabilistic Assessment

U.S. Environmental Protection Agency
New York, NY
April 21-22, 1998

Charge to Experts/Discussion Issues

This workshop is being held to discuss issues associated with the selection of probability distributions to represent exposure factors in a probabilistic risk assessment. The workshop discussions will focus on generic technical issues applicable to any exposure data. It is not the intent of this workshop to formulate decisions specific to any particular exposure factors. Rather, the goal of the workshop is to capture a discussion of generic issues that will be informative to Agency assessors working with a variety of exposure data.

On May 15, 1997, the U.S. Environmental Protection Agency (EPA) Deputy Administrator signed the Agency's "Policy for Use of Probabilistic Analysis in Risk Assessment." This policy establishes the Agency's position that "such probabilistic analysis techniques as Monte Carlo Analysis, given adequate supporting data and credible assumptions, can be viable statistical tools for analyzing variability and uncertainty in risk assessments." The policy also identifies several implementation activities designed to assist Agency assessors with their review and preparation of probabilistic assessments. These activities include a commitment by the EPA Risk Assessment Forum (RAF) to organize workshops or colloquia to facilitate the development of distributions for exposure factors.

In the summer of 1997, a technical panel, convened under the auspices of the RAF, began work on a framework for selecting input distributions for use in Monte Carlo analyses. The framework emphasized parametric methods and was organized around three fundamental activities: selecting candidate theoretical distributions, estimating the parameters of the candidate distributions, and evaluating the quality of the fit of the candidate distributions. In September of 1997, input on the framework was sought from a 12 member panel of experts from outside of the EPA. The recommendations of this panel include:

- expanding the framework's discussion of exploratory data analysis and graphical methods for assessing the quality of fit,
- discussing distinctions between variability and uncertainty and their implications,
- discussing empirical distributions and bootstrapping,
- discussing correlation and its implications,
- making the framework available to the risk assessment community as soon as possible.

Subsequent to receiving this input, some changes were made to the framework and it was applied to selecting distributions for three exposure factors: water intake per body weight, inhalation rate, and residence time. The results of this work are presented in the attached report entitled “Development of Statistical Distributions for Exposure Factors.”

Applying the framework to the three exposure factors highlighted several issues. These issues resolved into two broad categories: issues associated with the representativeness of the data, and issues associated with using the empirical distribution function (or resampling techniques) versus using a theoretical parametric distribution function. Summaries for these issues are presented in the attached issue papers. These issues will be the focal point for discussions during this workshop. The following questions are intended to help structure and guide these discussions. In addressing these questions, workshop participants are asked to consider: what do we know today that can be applied to answering the question or providing additional guidance on the topic; what short term studies (e.g., numerical experiments) could be conducted to answer the question or provide additional guidance; and what longer term research may be needed to answer the question or provide additional guidance.

Representativeness (Issues Paper #1)

1) The Issue Paper

Checklists I through IV in the issue paper present a framework for characterizing and evaluating the representativeness of exposure data. This framework is organized into three broad sets of questions: questions related to differences in populations, questions related to differences in spatial coverage and scale, and questions related to differences in temporal scale. Do these issues cover the most important considerations for representativeness? Are the lists of questions associated with each issue complete? If not, what questions should be added?

In a tiered approach to risk assessment (e.g., a progression from simpler screening level assessments to more complex assessments), how might the framework be tailored to each tier? For example, is there a subset of questions that adequately addresses our concerns about representativeness for a screening level risk assessment?

2) Sensitivity

The framework asks how important are (or how sensitive is the analysis to) population, spatial, and temporal differences between the sample (for which you have the data) and the population of interest. For example, to what extent do these differences affect our estimates of the mean and variance of the population and what is the magnitude and direction of these effects?

What guidance can be provided to help answer these questions? What sources of information exist to help with these questions? Having answered these questions what are the implications for the use of the data (e.g., use of the data may be restricted to screening level assessments in certain

circumstances)? What differences could be considered critical (i.e., what differences could lead to the conclusion that the assessment can't be done without the collection of additional information)?

3) Adjustments

The framework asks, is there a reasonable way of adjusting or extrapolating from the sample (for which you have data) to the population of interest in terms of the population, spatial, and temporal characteristics? If so, what methods should be used? Is there adequate information available to implement these methods?

What guidance can be provided to help answer these questions? Can exemplary methods for making adjustments be proposed? What sources of information exist to help with these questions? What research could address some of these issues?

Section 5 of the issue paper on representativeness describes methods for adjustments to account for differences in population and temporal scales. What other methods exist? What methods are available for spatial scales? Are there short-term studies that can be done to develop these methods further? Are there data available to develop these methods further? Are there numerical experiments (e.g., simulations) that can be done to explore these methods further?

Empirical Distribution Functions and Resampling Versus Parametric Distributions **(Issues Paper #2)**

1) Selecting the EDF or PDF

What are the primary considerations for assessors in choosing between the use of theoretical parametric distribution functions (PDFs) and empirical distribution functions (EDFs) to represent an exposure factor? Do the advantages of one method significantly outweigh the advantages of the other? Is the choice inherently one of preference? Are there situations in which one method is clearly preferred over the other? Are there circumstances in which either method of representation should not be used?

2) Goodness of Fit

On what basis should it be decided whether or not a data set is adequately represented by a fitted analytic distribution? What role should the goodness-of-fit test statistic play (e.g., chi-square, Kolmogorov-Smirnov, Anderson-Darling, Cramer-von Mises, etc.)? How should the level of significance, i.e., p-value, of the goodness of fit statistic be chosen? What are the implications or consequences for exposure assessors when acceptance/rejection is dependent on the goodness of fit statistic chosen and an arbitrary level of statistical significance? What role should graphical examination of the quality of fit play in the decision as to whether a fit is acceptable or not?

When the only data readily available are summary statistics (e.g., selected percentiles, mean, and variance), are fits to analytic distributions based on those summary statistics acceptable? Should any limitations or restrictions be placed in these situations?

When the better known theoretical distributions (e.g., lognormal, gamma, Weibull, log-logistic, etc.) cannot provide an acceptable fit to a particular set of data, is there value in testing the fit of the more flexible generalized distributions (e.g., the generalized gamma and generalized F distributions) even though they are considerably more complicated and difficult to work with?

3) Uncertainty

Are there preferred methods for assessing uncertainty in the fitted parameters (e.g., methods based on maximum likelihood and asymptotic normality, bootstrapping, etc.)?